

CLAIMS

What is claimed is:

1 1. A method for performing a frequent itemset operation, the method comprising the
2 steps of:
3 dynamically selecting which occurrence counting technique to use from a plurality of
4 available occurrence counting techniques; and
5 during said frequent itemset operation, using said selected occurrence counting
6 technique to count occurrences of at least one combination to determine
7 whether said at least one combination satisfies frequency criteria associated
8 with said frequent itemset operation.

1 2. The method of Claim 1 wherein:
2 the frequent itemset operation is performed in a plurality of phases, wherein each
3 phase is associated with combinations that have a particular number of items;
4 the step of dynamically selecting includes dynamically selecting which occurrence
5 counting technique to use for at least one phase of said plurality of phases; and
6 the step of using includes using said selected occurrence counting technique to
7 determine whether candidate combinations for said at least one phase satisfy
8 said frequency criteria.

1 3. The method of Claim 2 wherein:
2 said at least one phase is a phase during which combinations having N items are
3 processed;
4 a first occurrence counting technique is selected for said phase of said frequent
5 itemset operation;

6 the method includes dynamically selecting a second occurrence counting technique in
7 the phase of a subsequent frequent itemset operation during which
8 combinations having N items are processed; and
9 the second first occurrence counting technique is different from said second
10 occurrence counting technique.

1 4. The method of Claim 1 wherein the step of dynamically selecting which occurrence
2 counting technique includes generating cost estimates for each occurrence counting
3 technique of said plurality of available occurrence counting techniques, and selecting the
4 occurrence counting technique that has the lowest estimated cost.

1 5. The method of Claim 4 wherein the step of generating cost estimates includes
2 generating a cost estimate for at least one of said available occurrence counting techniques
3 based on an estimated I/O cost of using the occurrence counting technique.

1 6. The method of Claim 4 wherein the step of generating cost estimates includes
2 generating a cost estimate for at least one of said available occurrence counting techniques
3 based on an estimated CPU cost of using the occurrence counting technique.

1 7. The method of Claim 4 wherein the step of generating cost estimates includes
2 generating a cost estimate for at least one of said available occurrence counting techniques
3 based on a total cost that includes a plurality of weighted constituent costs.

1 8. The method of Claim 1 wherein the plurality of available occurrence counting
2 techniques include a bitmap intersection technique and a prefix tree technique.

1 9. The method of Claim 1 wherein the step of dynamically selecting includes
2 dynamically selecting based on conditions existing in a computing environment in which the
3 frequent itemset operation is to be performed.

1 10. The method of Claim 9 wherein the conditions include one or more of
2 workload of a computer system executing the frequent itemset operation; and
3 resources available on said computer system.

1 11. The method of Claim 2 further comprising the step of determining that a particular
2 occurrence counting technique will not be considered during any phase of the frequent
3 itemset operation, and performing the frequent itemset operation without performing startup
4 operations for said particular occurrence counting technique.

1 12. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to perform the
3 method recited in Claim 1.

1 13. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to perform the
3 method recited in Claim 2.

1 14. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to perform the
3 method recited in Claim 3.

1 15. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to perform the
3 method recited in Claim 4.

1 16. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to perform the
3 method recited in Claim 5.

1 17. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to perform the
3 method recited in Claim 6.

1 18. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to perform the
3 method recited in Claim 7.

1 19. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to perform the
3 method recited in Claim 8.

1 20. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to perform the
3 method recited in Claim 9.

1 21. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to perform the
3 method recited in Claim 10.

1 22. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to perform the
3 method recited in Claim 11.